GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2004, 14:59:54 ; Search time 3 Seconds
(without alignments)
4.215 Million cell updates/sec

Title:
Perfect score: 798
Sequence: 1 MEIIKKFVALGLLSAVLSSS............VVTAPNKIQLDMVSWGVGSIC 148

SEQ809

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 497 seqs, 85440 residues

Total number of hits satisfying chosen parameters: 497

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : US0487032A.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	DB	ID	Description
1	775	97.1	170	1	US-08-487-032A-764	Sequence 764, App
2	125.5	15.7	80	1	US-08-487-032A-808	Sequence 808, App
3	93.5	11.7	157	1	US-08-487-032A-403	Sequence 403, App
4	67.5	8.5	325	1	US-08-487-032A-833	Sequence 833, App
5	64	8.0	264	1	US-08-487-032A-498	Sequence 498, App
6	55.5	7.0	560	1	US-08-487-032A-708	Sequence 708, App
7	52.5	6.6	343	1	US-08-487-032A-632	Sequence 632, App
8	49	6.1	127	1	US-08-487-032A-595	Sequence 595, App
9	49	6.1	155	1	US-08-487-032A-777	Sequence 777, App
10	49	6.1	1178	1	US-08-487-032A-473	Sequence 473, App
11	46	5.8	214	1	US-08-487-032A-663	Sequence 663, App
12	46	5.8	325	1	US-08-487-032A-671	Sequence 671, App
13	44.5	5.6	199	1	US-08-487-032A-651	Sequence 651, App
14	44.5	5.6	237	1	US-08-487-032A-482	Sequence 482, App
15	44	5.5	308	1	US-08-487-032A-573	Sequence 573, App
16	44	5.5	421	1	US-08-487-032A-479	Sequence 479, App
17	43.5	5.5	351	1	US-08-487-032A-639	Sequence 639, App
18	43	5.4	210	1	US-08-487-032A-725	Sequence 725, App
19	43	5.4	496	1	US-08-487-032A-629	Sequence 629, App
20	42.5	5.3	187	1	US-08-487-032A-445	Sequence 445, App
21	42.5	5.3	237	1	US-08-487-032A-625	Sequence 625, App
22	42	5.3	164	1	US-08-487-032A-817	Sequence 817, App
23	42	5.3	422	1	US-08-487-032A-548	Sequence 548, App
24	41.5	5.2	509	1	US-08-487-032A-384	Sequence 384, App
25	41	5.1	121	1	US-08-487-032A-698	Sequence 698, App
26	41	5.1	162	1	US-08-487-032A-612	Sequence 612, App
27	41	5.1	702	1	US-08-487-032A-585	Sequence 585, App
28	40.5	5.1	240	1	US-08-487-032A-854	Sequence 854, App
29	40	5.0	133	1	US-08-487-032A-386	Sequence 386, App
30	40	5.0	179	1	US-08-487-032A-853	Sequence 853, App
31	40	5.0	200	1	US-08-487-032A-751	Sequence 751, App
32	40	5.0	313	1	US-08-487-032A-535	Sequence 535, App
33	39.5	4.9	198	1	US-08-487-032A-626	Sequence 626, App
34	39.5	4.9	231	1	US-08-487-032A-560	Sequence 560, App
35	39.5	4.9	233	1	US-08-487-032A-503	Sequence 503, App
36	39.5	4.9	309	1	US-08-487-032A-816	Sequence 816, App
37	39.5	4.9	339	1	US-08-487-032A-574	Sequence 574, App
38	39	4.9	173	1	US-08-487-032A-786	Sequence 786, App
39	39	4.9	193	1	US-08-487-032A-413	Sequence 413, App
40	39	4.9	193	1	US-08-487-032A-571	Sequence 571, App
41	39	4.9	278	1	US-08-487-032A-404	Sequence 404, App
42	39	4.9	286	1	US-08-487-032A-703	Sequence 703, App
43	39	4.9	341	1	US-08-487-032A-601	Sequence 601, App
44	38.5	4.8	213	1	US-08-487-032A-490	Sequence 490, App
45	38.5	4.8	237	1	US-08-487-032A-519	Sequence 519, App

ALIGNMENTS

RESULT 1
US-08-487-032A-764
Sequence 764, Application US/08487032A

; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 880
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,032A
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 764:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
US-08-487-032A-764

Query Match 97.1%; Score 775; DB 1; Length 170;
Best Local Similarity 98.6%; Pred. No. 3.4e-70;
Matches 144; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy    1 MEIIKKFVALGLLSAVLSSSLLAEGDGYYIGTNYQLGQARLNSNIYNTGDCTGSVVGCPP 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MEXXKKFVALGLLSAVLSSSLLAEGDGVYIGTNYQLGQARLNSNIYNTGDCTGSVVGCPP 60

Qy   61 GLTANKHPGGTNINWHSKYANGALNGFGLNVGYKKFQFKSLDMTSKWFGFRVYGLFDY 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GLTANKHPGGTNINWHSKYANGALNGFGLNVGYKKFQFKSLDMTSKWFGFRVYGLFDY 120

Qy  121 GHADLGKQVTAPNKIQLDMVSWGVGS 146
         ||||||||||||||||||||||||||
Db  121 GLTANKHPGGTNINWHSKYANGALNGFGLNVGYKKFQFKSLDMTSKWFGFRVYGLFDY 146
```

```
;  TELEPHONE: (617)227-7400
;  TELEFAX: (617)227-5941
; SEQUENCE CHARACTERISTICS:
;  LENGTH: 325 amino acids
;  TYPE: amino acid
;  TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
;  ORGANISM: Helicobacter pylori
US-08-487-032A-833

Query Match   8.5%;  Score 67.5;  DB 1;  Length 325;
Best Local Similarity 22.0%;  Pred. No. 0.12;
Matches 27;  Conservative 19;  Mismatches 50;  Indels 27;  Gaps 4;

Qy   23 AEGDGVVIGTNYQLGQARLNSNIYNTGDCTGSVVGCPPGLTANKHNPGGTNINWHSKYAN 82
        :|::|:::|||::|:|::||:|||                     |||:
Db  186 SENGGFNQNSKNIDKSLVNNMLFNTYLRTGGMM------------NKHFGIQAQVNW---- 231

Qy   83 GALNGFGL--------NVGYKKFFQFKSLDMTSKWFGFRVYGLFDYGHADLGKQVYAPN 133
        ||            :||::  |  |||:|||
Db  232 --LKGQGFRYNSPTDIQNYMLDSLYQINDSNKITAFQYYSYFLTDPG--SLGIAAYNQN 287

Qy  134 KIQ 136
        :|:
Db  288 RFQ 290
```

RESULT 5
US-08-487-032A-498
Sequence 498, Application US/08470032A
GENERAL INFORMATION:
 APPLICANT: DOUGLAS SMITH
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
 TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI FOR
 TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 880
 CORRESPONDENCE ADDRESS:
  ADDRESSEE: LAHIVE & COCKFIELD
  STREET: 60 State Street
  CITY: Boston
  STATE: Massachusetts
  COUNTRY: USA
  ZIP: 02109-1875
 COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/487,032A
  FILING DATE: 07-JUNE-1995
 ATTORNEY/AGENT INFORMATION:
  NAME: Mandragouras, Amy E.
  REGISTRATION NUMBER: 36,207
  REFERENCE/DOCKET NUMBER: GTN-001
 TELECOMMUNICATION INFORMATION:
  TELEPHONE: (617)227-7400
  TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 498:
 SEQUENCE CHARACTERISTICS:
  LENGTH: 264 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
  ORGANISM: Helicobacter pylori
 FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 1...264
  OTHER INFORMATION: /note= "vacuolating cytotoxin Hpylori"
US-08-487-032A-498

Query Match   8.0%; Score 64; DB 1; Length 264;
Best Local Similarity 29.4%; Pred. No. 0.2;
Matches 32; Conservative 14; Mismatches 27; Indels 36; Gaps 9;

```
Qy   13 LSAVLSSSLL--ABGDGV--YIGTN---YQLGQARLNSNIYNTGDCTGSVVGCPPGLTAN 65
        :|||:|||  ||||:||  :||::    |||||||||||||||||||||||||||||||
Db  183 VTARLGRDLLIKAKGDNVVRFVGENTLLYRKGB------IFNI---FASVI------- 224

Qy   66 KHNPGGTNINWHSKYANGALNGFGLNVGYKKFFQFKSLDMTSKWFGFRV 114
        ||| |:||||||||||:||        |:||:||||||||||:
Db  225 ---TGGEMHLWRLMYVNA----GVGLKMG----LQYQDLNITGN-VGMRV 262
```

RESULT 6
US-08-487-032A-708
Sequence 708, Application US/08470032A
GENERAL INFORMATION:
 APPLICANT: DOUGLAS SMITH
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
 TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI FOR
 TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 880
 CORRESPONDENCE ADDRESS:
  ADDRESSEE: LAHIVE & COCKFIELD
  STREET: 60 State Street
  CITY: Boston
  STATE: Massachusetts
  COUNTRY: USA
  ZIP: 02109-1875
 COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/487,032A
  FILING DATE: 07-JUNE-1995
 ATTORNEY/AGENT INFORMATION:
  NAME: Mandragouras, Amy E.
  REGISTRATION NUMBER: 36,207
  REFERENCE/DOCKET NUMBER: GTN-001
 TELECOMMUNICATION INFORMATION:
  TELEPHONE: (617)227-7400
  TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 708:
 SEQUENCE CHARACTERISTICS:
  LENGTH: 560 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
  ORGANISM: Helicobacter pylori
 FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 1...560
  OTHER INFORMATION: /note= "vacuolating cytotoxin"
US-08-487-032A-708

Query Match   7.0%; Score 55.5; DB 1; Length 560;
Best Local Similarity 25.9%; Pred. No. 3.2;
Matches 15; Conservative 9; Mismatches 23; Indels 11; Gaps 1;

```
Qy   66 KHNPGGTNINWHSKYANGALNGFGLNVGYKKFFQFKSLDMTSKWFGFRVYGLFDYGHA 123
        |:|               ||:||||||:||||:|       ||:|:
Db  289 KNNWATGVGGASFISGGTVXTLVGINXGYDRFIK--------GVIVGGYAAYGYS 335
```

RESULT 7
US-08-487-032A-632

; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
US-08-487-032A-777

Qy 101 KSLDMTSKW----------------FGFRVYGLFDYGHAD-------------LGKQ--- 128
          :|:|| | : |||.||: |||
Db 14 RSKALTRSWGINFSLKNTLAYAPWGFFDYAHANSIKLKNPNYNSEAAQVASQILGKQEIN 73

Qy 129 ----VVAPNKIQLDMVSWG 143
            :::|:::|
Db 74 RLTNIADPRTFEPNMLTYG 92

RESULT 10
US-08-487-032A-473
; Sequence 473, Application US/0847032A
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 880
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,032A
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 473:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1178 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...1178
; OTHER INFORMATION: /note= "cytotoxicity associated...immunodominant antigen [H.

Query Match
Best Local Similarity 20.3%; Score 49; DB 1; Length 155;
Matches 16; Conservative 10; Mismatches 17; Indels 36; Gaps 3;

Query Match
Best Local Similarity 44.4%; Score 49; DB 1; Length 1178;
Matches 16; Conservative 1; Mismatches 31; Indels 15; Gaps 4;

Qy 71 GTNINWHSKY--ANGALNGF--GLNVGYKKFFQFKS 102
         ||:|:| | :|: | |:|| | |||
Db 745 GINPEWISKVENLNAALNEFKNGKNKDFSKVTQAKS 780

RESULT 11

US-08-487-032A-663
; Sequence 663, Application US/0847032A
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 880
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,032A
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 663:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori

Query Match
Best Local Similarity 46.2%; Score 46; DB 1; Length 214;
Matches 12; Conservative 2; Mismatches 6; Indels 6; Gaps 1;

Qy 89 GLNV-----GYKKFFQFKSLDMTSK 108
         |||| |||:|||:| |:|
Db 129 GLNVELQINPEVKDFFTFKSISTTNK 154

RESULT 12
US-08-487-032A-671
; Sequence 671, Application US/0847032A
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 880
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,032A

```
Qy      65 NKHNPGGTNINWHSKYANG----- ALNG------FGLNVGYKFFQFKSLDMTSKWFGPRVY 115
            ::|::|:|:|:|:|        |:|:     :  :|||::|:|:||:| ::|:
Db      35 NTLSPIWSEINTYGIFAAGYLARPLGGIVMAHFGDKFGRKNMFMLSILLMVIPTFALALM 94

Qy     116 GLFDYGHADLGKQVYAPNKIQLDMVSWGVGSI 147
            |:|      ::|||::|:|| :
Db      95 PTFN-----------------DLVGFGVDSM 108

RESULT 15
US-08-487-032A-573
; Sequence 573, Application US/08487032A
; GENERAL INFORMATION:
;  APPLICANT: DOUGLAS SMITH
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
;  TITLE OF INVENTION: RELATING TO HELICOBACTER PYLORI FOR
;  TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
;  NUMBER OF SEQUENCES: 880
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE: LAHIVE & COCKFIELD
;   STREET: 60 State Street
;   CITY: Boston
;   STATE: Massachusetts
;   COUNTRY: USA
;   ZIP: 02109-1875
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/487,032A
;   FILING DATE: 07-JUNE-1995
;  ATTORNEY/AGENT INFORMATION:
;   NAME: Mandragouras, Amy E.
;   REGISTRATION NUMBER: 36,207
;  REFERENCE/DOCKET NUMBER: GTN-001
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (617)227-7400
;   TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 573:
;  SEQUENCE CHARACTERISTICS:
;   LENGTH: 308 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;  MOLECULE TYPE: protein
;  HYPOTHETICAL: YES
;  ORIGINAL SOURCE:
;   ORGANISM: Helicobacter pylori
;  FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: 1...308
;   OTHER INFORMATION: /note= "ToxR-activated (tagE) gene
;   OTHER INFORMATION: [vibrio cholerae] (inner membrane)"
US-08-487-032A-573

Query Match             5.5%;  Score 44;  DB 1;  Length 308;
Best Local Similarity   20.2%; Pred. No. 20;
Matches  24;  Conservative 19; Mismatches  42;  Indels  34; Gaps  5;

Qy      15 AVLSSLLAEGDGV--YIGTNYQLGQARL--------------NSNIYNTGDCT----------GS 54
            |:::|:|:|| |:  ::||::   |                 :::  :| |          ::
Db     172 APLNTPVYASADGIVDFVKTRSNAGYGNLVRIEHAFGFSSIYTHLDHVNVQPKSFIQKGQ 231

Qy      55 VVGCPPGLTANKHNPGGTNINWH------------------SKYANGALNGFGLNVGYKKFFQFKSL 103
            :: |:: :|:| ::|:     ::                 :::|||:: :::|
Db     232 LI----GYSGKGGNSGGEKLHYEVRFLGKILDAEKFLAWDLDHFQSALEENKFIEWKNL 286
```